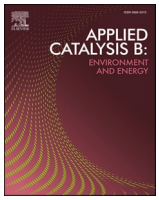
Applied Catalysis B: Environment and Energy 353 (2024) 124070

Contents lists available at ScienceDirect

Applied Catalysis B: Environment and Energy

journal homepage: www.elsevier.com/locate/apcatb



Predicting hydrogenolysis reaction barriers of large hydrocarbons on metal surfaces using machine learning: Implications for polymer deconstruction

Xue Zong [a,b], Tianjun Xie [b], Dionisios G. Vlachos [a,b,c,*]

[a] Department of Chemical and Biomolecular Engineering, 150 Academy St., University of Delaware, Newark, DE 19716, United States
[b] Delaware Energy Institute (DEI), 221 Academy St., University of Delaware, Newark, DE 19716, United States
[c] Center for Plastics Innovation, University of Delaware, 221 Academy St., University of Delaware, Newark, DE 19716, United States

ARTICLE INFO

Keywords:
Hydrogenolysis
Activation energies
Scaling relationships
Machine learning
Catalyst screening
Plastic (up)recycling

ABSTRACT

Calculating activation energies of chemical reactions for large reaction networks is computationally demanding. Traditional Brønsted-Evans-Polanyi (BEP) relationships are useful in this regard but are prone to substantial estimation errors. Here, we explore machine learning models to predict activation energies for hydrogenolysis reactions on transition metals, including Ru, Pt, and Rh. The curated dataset includes 380 DFT-calculated activation energies of C-C and C-H scission reactions ranging from C1 to C6 species. The mean absolute error of traditional BEPs is 0.30 eV, and published BEPs on smaller hydrocarbons give even larger errors. In comparison, an XGBoost model achieves a mean absolute error of 0.19 eV for the test set without needing additional DFT-calculated energies. The model selects the homologous series among reactions and the electronegativity and atomic mass of the metal atoms as key features. The model appears transferable across metals. This approach can provide accurate and rapid estimation of activation energies of large reaction networks, such as those in plastics recycling, to accelerate catalyst screening. We showcase the impact of hydrocarbon chain length and branching on the barriers, discuss implications for the depolymerization rate of isotactic polypropylene and low vs. high-density polyethylene, and demonstrate the feasibility of hydrogenolysis catalyst screening.

1. Introduction

Plastic waste accumulation in oceans and soil has become an acute environmental crisis [1]. This issue intensified during the recent pandemic [2], marked by a steep increase in plastic use. This surge increased reliance on incineration, contributing to extensive CO_2 emissions [3], and landfilling of consumer packaging, foams, films, and personal protective equipment [4]. A significant portion of waste comprises polyolefins, like polyethylene and polypropylene [5], whose recycling rates are low. Chemical recycling [6–8], specifically through hydrogenolysis [9,10], is a promising approach to convert polyolefins into valuable products. It involves breaking down C-C bonds in polyolefins using hydrogen and a noble metal catalyst, such as Ru and Pt, supported on carbon or metal oxides [11–15]. Developing more efficient catalysts could greatly enhance recycling rates, but the complex reaction networks make it challenging to design materials computationally.

Microkinetic modeling (MKM) has become an effective approach for evaluating catalyst performance and screening potential catalysts [16–22] but requires accurate MKM parameters, particularly activation energies. These are typically derived from density functional theory (DFT) calculations. However, calculating activation barriers of increasingly complex reactions is computationally demanding. To address this, linear correlations, such as the Brønsted-Evans-Polanyi (BEP) [23–25] relationships, have been developed that estimate reaction barriers by relating them to reaction energies, which are less expensive to calculate [26–28]. This approach estimates all elementary reaction barriers using a few DFT calculations. Despite their utility, BEP correlations are not universally applicable and often require reparameterization for different reaction types. Their predictive ability in extrapolation has not generally been tested, as most calculations employ small molecules.

To increase the accuracy of predicted activation energies of various reactions, machine learning (ML) and deep learning algorithms become more common [29–33]. Nørskov and co-workers [34] developed an ML model to predict reaction barriers of dissociation reactions, with a mean absolute error (MAE) of 0.25 eV, by having 7 additional descriptors besides the reaction energy. Goltl et al. [35] developed generalized BEP relationships (gBEPs) by adding features to describe the change in chemical identities in a reaction. They achieved an error of 0.23 eV for

* Corresponding author at: Department of Chemical and Biomolecular Engineering, 150 Academy St., University of Delaware, Newark, DE 19716, United States.
E-mail address: vlachos@udel.edu (D.G. Vlachos).

https://doi.org/10.1016/j.apcatb.2024.124070
Received 15 January 2024; Received in revised form 7 April 2024; Accepted 11 April 2024
Available online 12 April 2024

the test set. Their recent work [36] predicted activation energies with an MAE of 0.19 eV, utilizing their gBEPs without the reaction energy in the descriptor set. Despite promising results, sizeable datasets of heterogeneous catalysis are missing, and ML models with a few features, free of first-principles calculations, are needed.

In this study, we first generate a DFT dataset of activation barriers for C-C and C-H bond breaking reactions relevant to hydrogenolysis of the catalytic recycling of plastics. The dataset includes 380 DFT-calculated activation energies, 60 of which were published in our previous work [37] and the rest involve larger hydrocarbon species up to C_6 on Ru, Pt, and Rh surfaces. We develop BEP relationships for both C-C cracking and C-H breaking reactions. We observe that C-H breaking reactions exhibit better correlations than C-C cracking reactions, and BEP correlations on different metals remain similar. However, the error in predicting the energies of larger molecules, by using BEPs data of smaller molecules, is high. The error of BEPs including larger species is smaller but still relatively high. We explore various ML, non-BEP-related models for predicting activation energies with simple descriptors without relying on DFT calculations and find transferability across metals. Our model significantly reduces the computational cost of screening catalytic materials, thus accelerating the discovery in material space. Finally, we use the ML models to predict barriers of larger hydrocarbons and on various metals outside the training set to expose the potential for understanding polymer deconstruction paths and rates.

2. Methods

2.1. DFT calculations

All electronic energies and vibrational frequencies were calculated using the Vienna ab initio Simulation Package [38,39] (VASP 5.4.4) and the Atomic Simulation Environment [40] (ASE 3.17.0). We used the projector augmented-wave (PAW) method [41,42] to describe the electron-ion interactions and the Perdew-Burke-Ernzerhoff (PBE) [43] functional of the generalized gradient approximation (GGA) to obtain exchange-correlation energies. The D3 correction [44] was applied to account for dispersion interactions. The Kohn-Sham one-electron valence states were expanded in plane-wave basis sets with a cutoff energy up to 400 eV. Gaussian smearing with a smearing factor of 0.1 eV was employed, and all potential energies were extrapolated to 0 K. The convergence criterion for the self-consistent electronic minimization was set to 10^{-6} eV, and the Brillouin zone was sampled using $3\times3\times1$ k-point meshes constructed using the Monkhorst-Pack scheme [45]. The conjugate gradient algorithm with a force convergence threshold of 0.02 eV/Å was used for ionic relaxations.

All metal surfaces were modeled as four-layer slabs, and the atoms in the two topmost layers were allowed to relax in all directions. At least 15 Å of vacuum was added vertically to minimize interactions between successive slabs, and dipole moment corrections were applied to avoid periodic interactions in the normal direction. The reaction energy was calculated as the difference between the final state energy E_{FS} and initial state energy E_{IS}, while the activation energy was defined as the difference between the transition state (TS) energy E_{TS} and initial state energy E_{IS} [46]. All energies considered in this work were electronic energies.

The climbing image nudged elastic band (CI-NEB) method [47] along with the dimer method [48] were used to search for the TSs of reactions, and TSs were verified by vibrational frequency analysis showing the existence of only one imaginary mode. The NEB projected reaction path was dissected into intermediate images and optimized until the residual force was below 0.05 eV/Å. The TS determined by the CI-NEB method was used as the initial input to the dimer method. The VIST tools were used for the CI-NEB, dimer, and harmonic frequency calculations [49].

2.2. Dataset

We consider the barriers of 266 C-C cracking and 114 C-H dehydrogenation reactions of C_1 to C_6 alkanes for a total of 380 data points. These reactions occur in many catalytic systems, such as plastics hydrogenolysis [37]. 60 of the reaction data is taken from our previous work [37]. The complete list of reactions is shown in the Supporting Data file. We chose Ru (0001), Pt (111), and Rh (111) surfaces of metals commonly used in hydrogenolysis based on the lowest surface free energy facet of each metal. For C-C cracking reactions, 101 reactions were calculated on Ru surfaces, 108 reactions on Pt, and 57 reactions on Rh. For each metal, 12 species were branched C_4 and C_5 fragments. For C-H breaking reactions, 84 reactions were calculated on Ru and 30 reactions on Pt. The distribution of data for each homologous series (C-C vs. C-H reaction type) and hydrocarbon size are summarized in Figure S1.

2.3. Machine learning (ML) models

In developing ML models, we included all reaction data of the 380 DFT calculations. The scikit-learn package [50] (version 1.0.2) was used to train, cross-validate, and test all ML models. 80% of the data was randomly selected and used to train a model, while the remaining 20 % was used to test the model performance. K-fold cross-validation (k=10) was used to tune the model and avoid overfitting. The XGBoost regressor [51,52] (2.0.1) based on tree ensembles was used. SHAP values developed by Lee et al. [53] were calculated for feature importance analysis. The correlation matrix was constructed using the Seaborn [54] package.

To achieve optimal model performance, the hyperparameters were tuned using the Hyperopt [55] package based on Bayesian Optimization. Hyperopt allows automation and employs Bayesian optimization to efficiently search through a predefined hyperparameter space to find the optimal hyperparameters for a given ML algorithm. Tree-structured Parzen Estimators (TPE) [56] were used to assess the performance of ML models based on different hyperparameter configurations. The optimal hyperparameters for each model are summarized in Table S1.

3. Results and discussion

3.1. BEP relationships

The reactions constitute a sizeable dataset for dehydrogenated hydrocarbons on three transition metals. BEP relationships for lighter alkanes on Pt have been developed before [46,57,58]. Here we assess if these BEP relationships hold for large hydrocarbon C-C and C-H bond cleavage across three metals. Given data clustering for each reaction type, we also regressed all points to develop a universal correlation.

The correlations for each homologous series are depicted in Fig. 1 and the parameters are listed in Table 1. Correlation plots for each metal are shown in Figure S2 and S3. The reaction energies are correlated with the activation energies indicating the existence of BEP relationships. The BEPs are more accurate for C-H reactions with a MAE of 0.16 eV for all metals and 0.09 eV for Pt. In contrast, the MAE for C-C cracking reactions is higher at 0.22 eV. C-C cracking reactions have larger activation barriers than C-H breaking reactions for comparable reaction energies (Fig. 1c). The reaction barrier depends not only on thermodynamics, but also on the intrinsic bond polarity [59]. The C-C and C-H bonds possess different bond polarities [58]. On metal surfaces, the hydrocarbons prefer to dehydrogenate first until C-C cracking becomes facile. Among metals, BEPs on Pt are more accurate than other metals for both reaction types. To assess if the BEPs derived from small hydrocarbon species can estimate reaction barriers for large hydrocarbon species accurately, we applied the literature BEPs to our dataset. The estimation errors are summarized in Table S2. BEPs of this work give better estimation results than literature's, underscoring the need to develop BEPs for larger hydrocarbon species. Furthermore, literature BEPs on Pt are also more accurate than Ru, consistent with our findings.

The slope of a BEP correlation reveals the structure of the transition state; a slope near unity suggests a late transition state and one close to zero an early one [60]. The intercept provides insights into the inherent

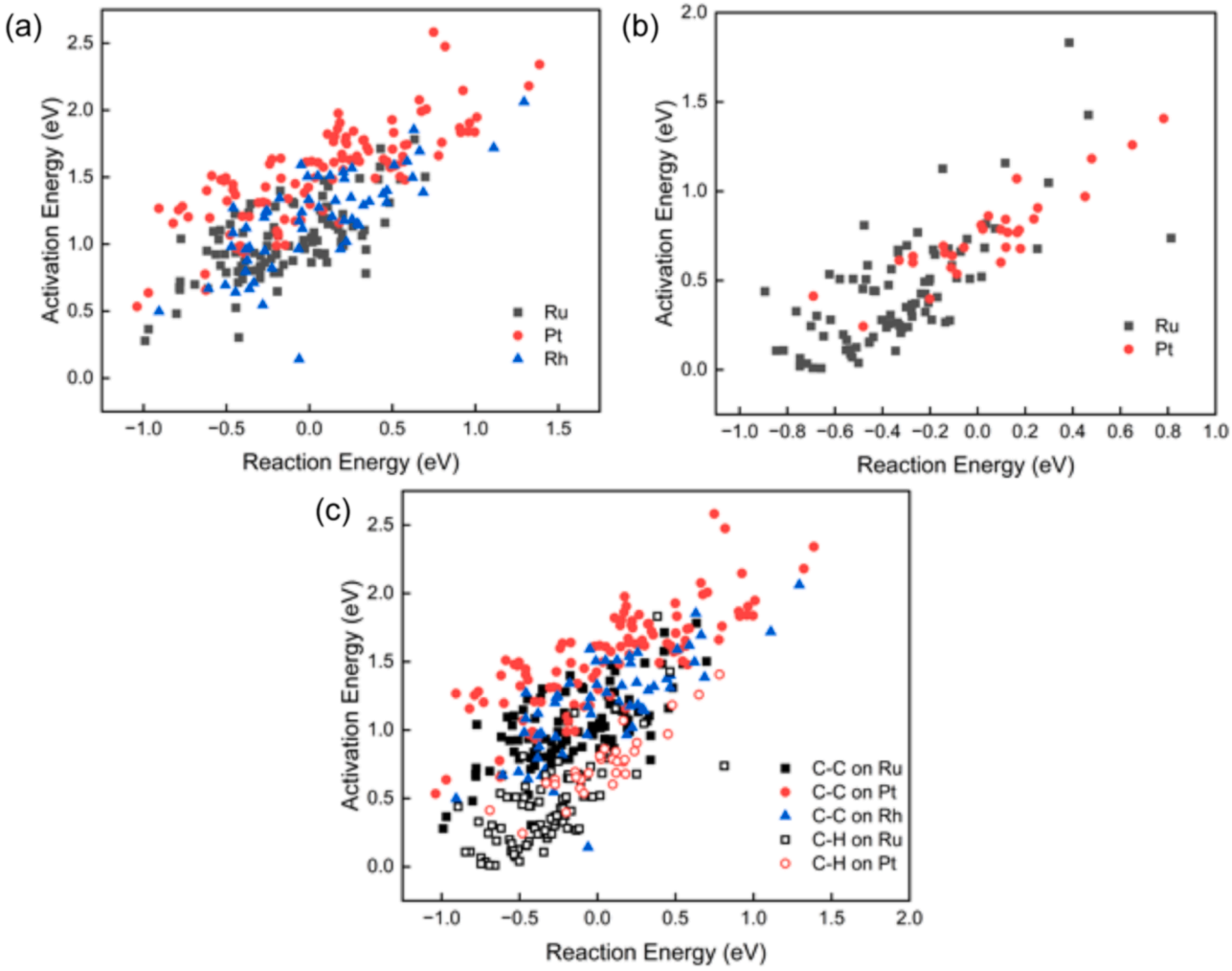


Fig. 1. BEP correlations on individual metals for (a) C-C reactions, (b) C-H reactions, (c) all data. Color code: black Ru; red Pt; blue Rh. Correlation parameters in Table 1.

Table 1
Summary of BEP relations of C-H and C-C bond scissions, number of points, R^2, and mean absolute error (MAE) in eV. R^2 values are correlation coefficients. The 95 % confidence intervals for the slopes and intercepts are indicated.

Homologous Series	Metal	Slope (eV)	Intercept (eV)	Number of Data Points	R^2	MAE (eV)
C-C	Ru	0.53 ± 0.11	1.09 ± 0.04	101	0.47	0.16
	Pt	0.55 ± 0.08	1.48 ± 0.04	108	0.63	0.18
	Rh	0.63 ± 0.15	1.16 ± 0.06	57	0.56	0.18
	All	0.64 ± 0.07	1.27 ± 0.03	266	0.53	0.22
C-H	Ru	0.75 ± 0.16	0.68 ± 0.07	84	0.50	0.18
	Pt	0.71 ± 0.14	0.73 ± 0.04	30	0.80	0.09
	All	0.77 ± 0.11	0.69 ± 0.04	114	0.63	0.16
All	All	0.80 ± 0.09	1.10 ± 0.04	380	0.47	0.30

reactivity of a metal [46]. Table 1 indicates that the BEP slopes are comparable across metals for the same homologous series, implying that the transition state geometries are similar. Furthermore, C-C reactions typically have moderately late to late transition states, whereas C-H reactions exhibit predominantly late transition states. These patterns align with their final state geometries of the dissociated fragments. Fig. 2 illustrates an example of the transition state configurations for C-C

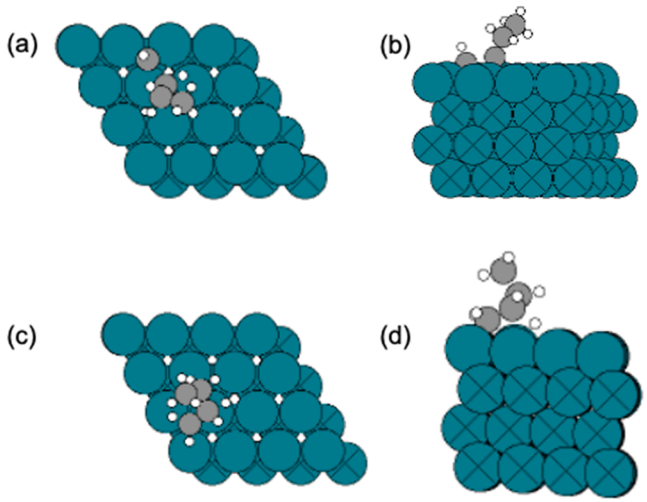


Fig. 2. Transition state configurations of C-C cracking and C-H breaking reactions. (a) C-C cracking top view, (b) C-C cracking side view, (c) C-H breaking top view, (d) C-H breaking side view.

cracking and C-H breaking reactions on Ru. C-H scission reactions result in a dehydrogenated species, where the carbon atom remains bound to the same metal atoms as before dissociation for large hydrocarbon fragments, and the hydrogen adsorbs on a nearby top or hollow site with minimal energy penalty, consistent with the literature [61]. Conversely, C-C cracking often occurs in species where carbon atoms are bound to neighboring sites on a metal surface. One or more cleaved fragments tend to be strongly bound on the surface [62], which requires diffusing to a next-nearest neighbor site to minimize the repulsive interactions that would otherwise occur, leading to increased bond lengths in the transition state. For C-C and C-H breaking reactions, the intercepts on Pt, indicative of the intrinsic activation barriers on metals, are slightly larger than the other two metals. This finding is consistent with the activity trend observed in experiments. Specifically, Sinfelt and co-workers [63,64] and Morita and co-workers [65] reported that the hydrogenolysis activity of n-pentane and n-heptane on transition metals follows the order Ru > Rh > Pt, consistent with the activation barriers shown in Fig. 1a.

Our analysis indicates that BEP relations for small hydrocarbons do not extrapolate well and for large hydrocarbon species are reasonably accurate. However, the overall MAE error of 0.30 eV when considering all data is rather large. For larger hydrocarbons, computing reaction energies takes considerable time and calculation of the final state energy is challenging due to many possible adsorption configurations of the products. To overcome limitations of traditional BEP relationships and provide a more accurate and rapid estimation of reaction energies and activation barriers, we turn to ML.

3.2. Featurization and correlations

Feature selection is crucial of a ML model. Inspired by previous studies [35,36,66], features describing the reactant, metal surfaces, and the reaction type, were selected (Table 2). No surface structure information was included, since only the (111) facet was considered. The first set of features includes the molecular weight of the reactant, the adsorbate valency of the two atoms where the bond is breaking, and the cis/trans structure of the reactant. The second includes metal surface information following our previous work [67], such as the atomic mass of the metal atom, electronegativity, first ionization potential, number of d electrons, and electron affinity. The reaction type is labelled as C-C breaking or C-H breaking. All features are physically relevant and easily determined without DFT calculations.

We analyzed the correlation among input features, reaction energies, and activation energies. The categorical variables were first converted to numerical values using one-hot encoding since no ordinal relationship exists. One-hot encoding converts each category value into a new binary column and assigns 0 or 1 to those columns indicating the presence of the categorical value. This transforms the categorical variable into a form that can be effectively used in regression algorithms. The resulting correlation matrix is shown in Fig. 3. The metal descriptors are strongly correlated with each other. The reaction energy is most intensively correlated with the activation energy, as expected, followed by the reaction type and the metal-related properties. The importance of including C-C cracking reaction as a descriptor arises from that the activation barriers of C-C cracking reactions cover a much wider range than C-H reactions, as shown in Fig. 1c.

Table 2
Features included in the machine learning (ML) model.

Category	Abbreviation	Variable Type	Feature Description
Reactant	Reactant_Mass	Numerical	Molecular weight of the reactant
	C1 valency	Numerical	Adsorbate valency of the first reacting atom that bonded to the metal surface
	C2 valency	Numerical	Adsorbate valency of the second reacting atom
	Cis/Trans	Categorical	Cis/trans structure of the reactant. Three types used: cis, trans, not applicable
Metal	Metal_Mass	Numerical	Atomic mass
	chi	Numerical	Electronegativity
	IP	Numerical	First ionization potential
	n_d	Numerical	Number of d electrons
	EA	Numerical	Electron affinity
Reaction	Reaction_Type	Categorical	Bond breaking type: C-C or C-H

3.3. Model performance

To predict the activation energies accurately, we exploited different ML regression algorithms and tuned their hyperparameters. We tested linear, kernel, bagging, and boosting models, including multivariate linear regression (MLR), support vector regression (SVR) [68], random forest regression (RFR) [69], and extreme gradient boosting regression (XGBoost) [51,52]. We prioritized simpler regression models to ensure interpretability. The MAE was chosen as the performance metric.

Our findings revealed that the simplest model, MLR, yielded a test set error of 0.24 eV. This indicates that even basic models can be effective when using well-selected features. Of all models tested, the XGBoost model outperformed the rest, achieving an MAE of 0.19 eV on the test set. The SVR and RFR models showed commendable performance, although the former required the longest training time due to its optimization process by solving a convex problem in a high-dimensional space. The XGBoost model performance is depicted in Fig. 3b and that of the other models in Figure S4 and Table S3. Errors for each reaction type and each individual metal are summarized in Table S4. Overall, the XGBoost model performance is good on all metals, with MAE no larger than 0.20 eV. Its MAE from all reactions is lower than a previous ML model [35] using DFT-calculated reaction energies as descriptors (a MAE of 0.23 eV) and comparable to a model [36] (a MAE of 0.19 eV) of 51 features for each reaction, while using fewer than ten easily estimated features.

To compare with traditional BEP correlations, we added the reaction energy as a feature in the XGBoost model. Upon tuning hyperparameters, the model's MAE test set was reduced from 0.19 eV to 0.16 eV (Fig. 4a), notably smaller than 0.30 eV of the traditional BEP relations. Even without DFT-calculated reaction energies, the model estimates activation energies accurately.

Our analysis has been using 80 % of the data points for training and 20 % for testing, a common split in the ML literature. To evaluate the XGBoost model's performance for different dataset sizes and ascertain

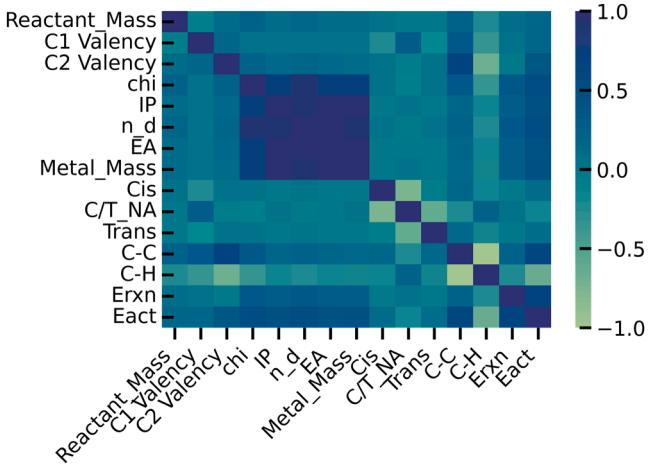


Fig. 3. Correlation between features, reaction energies, and activation energies. See Table 2 for a list of features. Deep blue (green) indicates a strong positive (inverse) correlation.

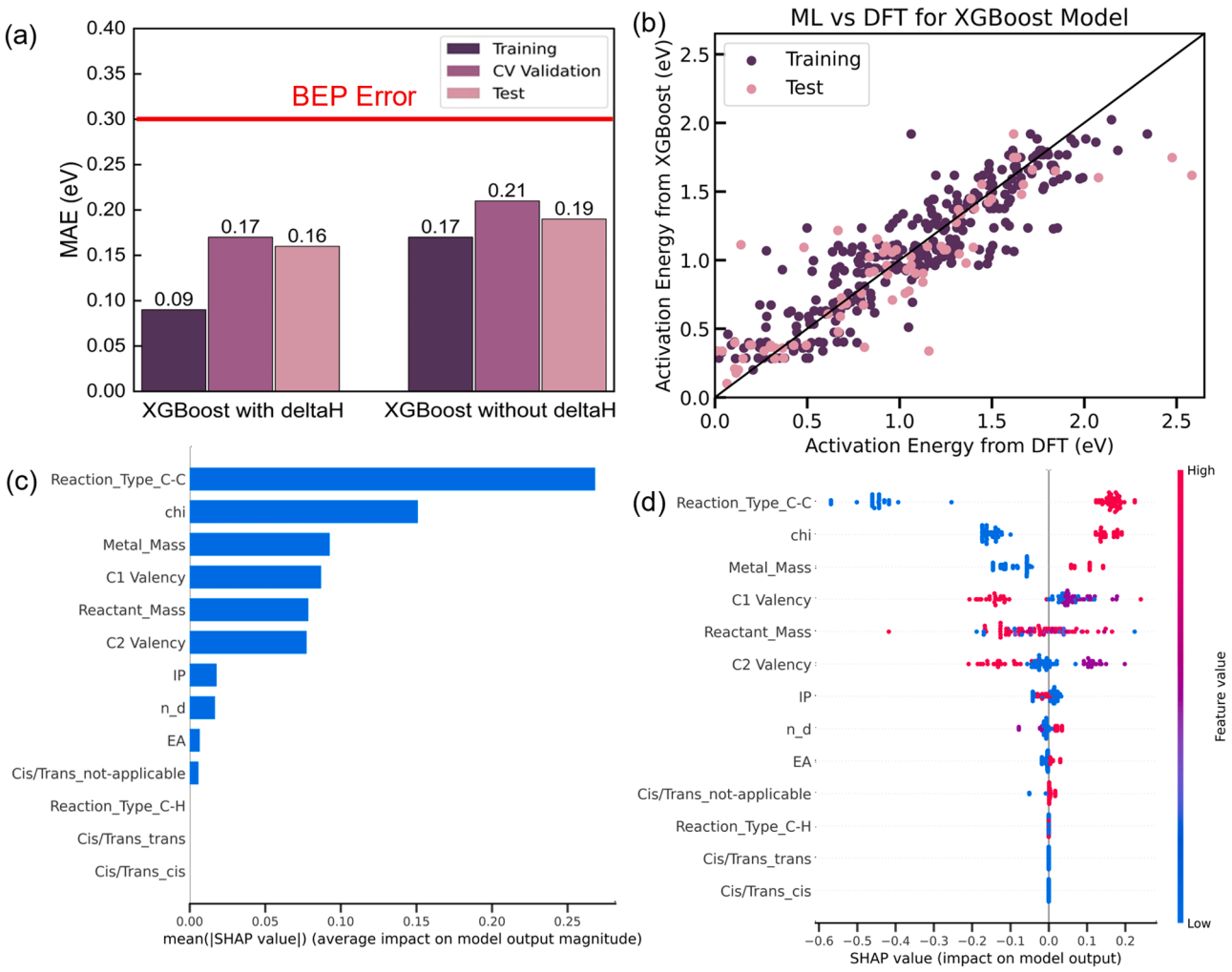


Fig. 4. ML model performance. (a) MAE of ML models with different feature sets. deltaH represents the reaction energy. (b) Activation energies from the XGBoost model vs. DFT-calculated values. Color code: purple points are the training set (80 % of data); pink points are the test set (20 % of data). (c) SHAP values of each feature for test data show the impact of each feature on the model output. (d) Feature importance based on mean SHAP values. See Table 2 for feature descriptions.

model convergence, we varied the test set size from 0 % to 90 %, adjusting the training set size accordingly in 10 % increments (Fig. 5). We observed a large test set error with a 90 % test set size. Interestingly, as we increased the training set size, the test set error decreased and eventually stabilized upon reaching 80 % of the training set, indicating no overfitting.

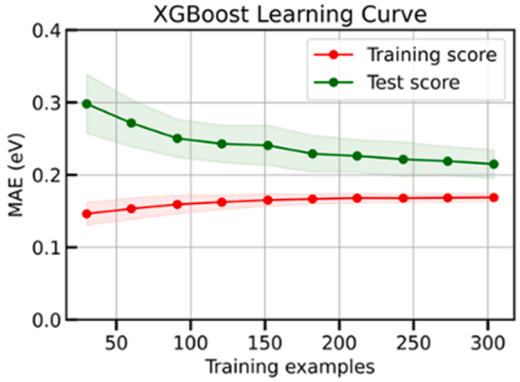


Fig. 5. Learning curve of the XGBoost model with different training and test set sizes.

3.4. Feature importance analysis

Identifying the key features that influence the activation energy is vital. For this purpose, we employed the SHAP (Shapley Additive exPlanations) method proposed by Lee et al. [53,70] to assess the feature importance of the XGBoost model. SHAP is an advanced AI technique grounded in game theory. It calculates Shapley values for interpreting ML outcomes. Fig. 4c and d display each feature's significance and impact on predicting activation energies. Fig. 4c orders the features by their average impact on model predictions. Notably, the reaction type, specifically the C-C cracking reaction, emerged as the most predictive feature, consistent with correlations in Fig. 3. The next two important features are the metal electronegativity and atomic weight, followed by the adsorbate valency and the molecular weight of the reactant. Fig. 4d shows the features in order of importance and how each feature affects the model output. For example, when the reaction_type_C-C has a higher value (redder in the plot), the model output tends to be larger, consistent with the C-C cracking reactions having higher activation barriers.

The cis/trans descriptor only has a minor effect. The model after excluding this feature gives a test set MAE of 0.20 eV, close to the original model. A plausible explanation could be the limited number of cis/trans structures in our dataset, potentially underrepresenting this feature's significance. Considering potential applications to broader chemistries, we opted to retain this feature in the model.

3.5. Transferability across metals

To evaluate the transferability of the XGBoost model, we applied it to datasets comprising various metal combinations and employed a method akin to cross-validation. We reserved data from one metal surface as a test set while training the model on the remaining two. The model's performance is illustrated in Fig. 6, and details are summarized in Table S5.

The best performance occurs when predicting Rh data using a model trained on Ru and Pt data, achieving a test set MAE of 0.22 eV. As depicted in Fig. 6a, the training data sufficiently encompasses the test data range. For the other two models, the overlap between training and test data is only partial, and the test dataset size increases significantly. For example, when predicting Ru data, the training dataset is nearly as large as the test dataset, leading to a higher test error. The highest error in these models of 0.28 eV is higher than 0.19 eV of the original model but competitive with traditional BEP correlations. This result suggests that the model could be useful for screening other transition metal surfaces not included in the training dataset. Enhancing the training dataset with more data for more metals would likely further improve its performance.

3.6. Implications for larger hydrocarbon hydrogenolysis

We demonstrate the utility of these ML models with select examples. First, we provide insights into the hydrogenolysis of larger hydrocarbons by considering C-C cracking of C_5 species on the three catalysts (Rh, Ru, and Pt). Cracking reactions of linear hydrocarbons exhibit distinct characteristics on different catalysts. On Rh, the cracking reaction with the lowest barrier (0.14 eV) occurs for CH_3CH-$CCHCH_3$* → CH_3CH* + $CCHCH_3$*. On Ru, the corresponding reaction is CH_3C-$CHCCH_3$* → CH_3C* + $CHCCH_3$*, with a barrier of 0.28 eV. Pt displays the lowest barrier for CH_3CH_2C-CCH_2* → CH_3CH_2C* + CCH_2*; yet, it has the highest barrier among the three catalysts at 0.53 eV. All three catalysts facilitate cracking in the middle of the carbon chain. Rh and Ru exhibit early-stage (less dehydrogenation) cracking on CH-C, whereas Pt induces late-stage (more dehydrogenated) cracking on C-C. The more dehydrogenated carbonaceous fragments may be harder to hydrogenate and desorb and could lead to coke or C-C scission to form excess methane. Clearly and expectedly, the 'weak links' from where C-C scission happens and, thus, the selectivity are metal catalyst specific.

Next, we consider branched hydrocarbons to understand how branching, for example in isotactic polypropylene (i-PP) and low vs. high density polyethylene (LDPE vs. HDPE), impacts the chemistry. We take isopentane, $CH_3CH(CH_3)$-CH_2CH_3, as an example. The cracking reactions follow a similar trend to linear hydrocarbons but with higher barriers. On Rh, $CH_3C(CH_3)$-CCH_3* → $CH_3C(CH_3)$* + CCH_3* has a barrier of 0.67 eV. On Ru, $CH_3CH(CH_3)$-CCH_3* → $CH_3CH(CH_3)$* + CCH_3* has a barrier of 0.80 eV. Pt shows the highest barrier at 1.49 eV for $CH_3C(CH_3)$-CCH_3* → $CH_3C(CH_3)$* + CCH_3*. All three catalysts induce cracking between the beta and gamma carbons upon sufficient dehydrogenation (removal of 2–3 Hs). The impact of metal appears to be more pronounced (larger variation in barrier) than for linear hydrocarbons, with Pt being the slowest catalyst among the three. Consistent with previous studies [71,72], branched species must lose additional H atoms via C-H activation to break the C-C bond.

Overall, like small (C_2, C_3) hydrocarbons, C-C cracking occurs after three (or four in the case of Pt) dehydrogenation steps for linear paraffins and 2–3 for branched paraffins. Branched hydrocarbons possess the lowest barrier at the branching points and generally exhibit higher reaction barriers than linear hydrocarbons despite the similar degree of dehydrogenation. This implies that i-PP and LDPE should depolymerize more slowly than HDPE at the atomic level (turnover frequency). Notably, linear hydrocarbons on Rh and Ru display prominent low cracking barriers, competitive with dehydrogenation steps. These low barriers also indicate that cracking rates increase with the carbon chain length. This is consistent with Hibbitts et al. [73] who reported that the calculated enthalpies in the C-C cleavage of C_2-C_{10} n-alkanes on Ir decrease with increasing alkane chain length due to attractive van der Waals interactions with surfaces.

Obviously, the analysis here is simplified; it considers only local quantum chemistry effects without kinetic models. It also leaves out the polymer conformations stemming from the macromolecular nature of polymers that depend on their molecular architecture, the fact that more C-C scissions must break for larger molecular weight polymers, etc. These aspects make it difficult to generalize the reactivity trends from small and medium-sized hydrocarbons to long-chain polymers. Another aspect not considered here is catalyst particle size effects. The particle size is expected to affect the adsorption of hydrocarbon species and the number of carbons on the surface. For example, recent molecular dynamics calculations on a Ru_{22} cluster on TiO_2 indicated a maximum number of carbons of ~5. Currently, all DFT calculations are on extended metal surfaces. The models introduced here provide a basis for understanding the depolymerization kinetics but refinements that account for conformations and particle size effects will be necessary.

3.7. Application of the ML model to catalyst screening

To ascertain the use of our ML model for catalyst screening beyond the three metals studied herein, we investigate the ethane hydrogenolysis reaction as an illustrative case study. Based on our previous work [37], for ethane hydrogenolysis, the C-C scission usually occurs from CHCH*. Therefore, we predicted the activation barriers for both CHCH* → 2CH* and CHCH* → CHC* + H* reactions across nine transition metals (Pd, Pt, Rh, Ru, Ag, Au, Cu, Ir, Ni). The results are shown in Fig. 7 (values summarized in Table S6). Among all metals, Ru shows the

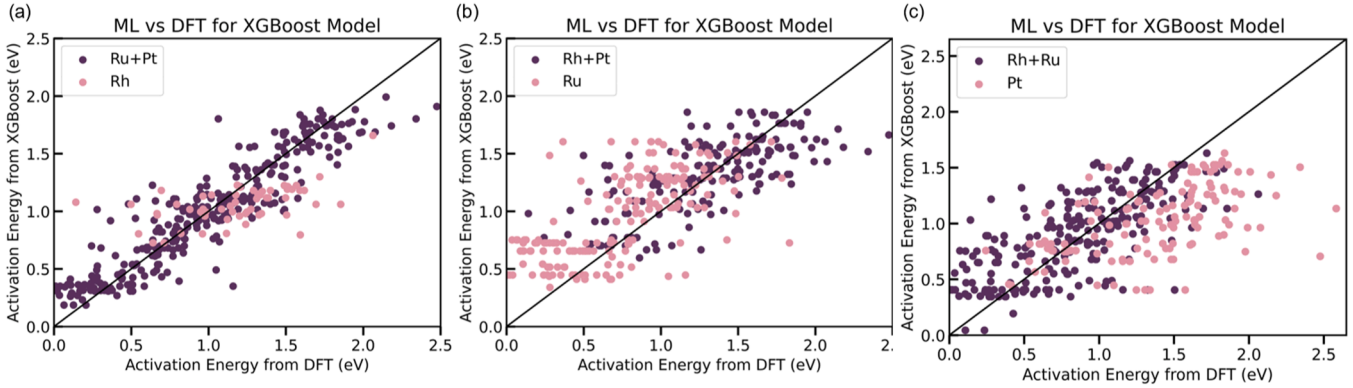


Fig. 6. Model transferability in predicting activation energies on (a) Rh, (b) Ru, and (c) Pt by training the model using data on the other metal surfaces. Color code: purple points represent the training set; pink points represent the test set.

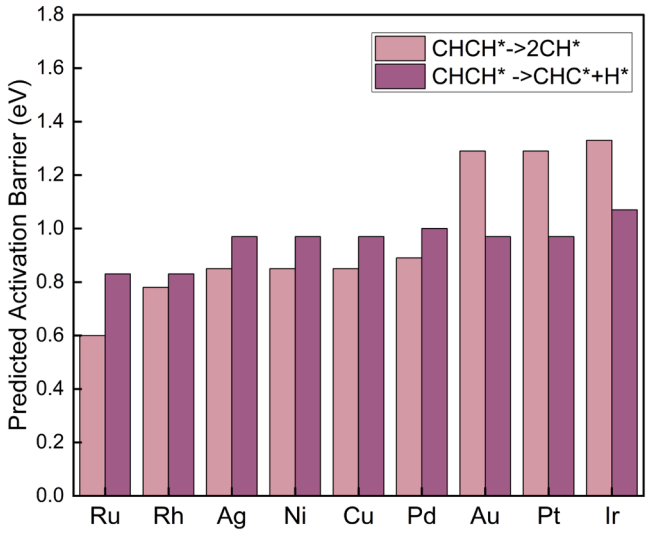


Fig. 7. Predicted activation barriers for C-C and C-H bond scission in CHCH* across metals.

lowest activation barrier for C-C cracking, followed by Rh. This order in barriers demonstrates that Ru- and Rh-based catalysts should be the fastest hydrogenolysis catalysts, consistent with literature [74,75]. Based on previously measured turnover rates from experiments, the activity of noble metals decreases in the order Ru > Rh > Ir > Pt. Notably, most of the metals demonstrated lower barrier for C-C bond scission, CHCH* → 2CH*, than dehydrogenation, CHCH* → CHC* + H*, signifying the prevalence of C-C scission from CHCH*. Considering a mean absolute error (MAE) of approximately 0.2 eV and the inclusion of only three metals in the training dataset, our ML model could be valuable to expeditiously screen potential metals in early stages of catalyst selection. Without kinetic models, our model captures the general activity trends and sheds light on potential catalysts but detailed kinetic models and parameter refinement via DFT are suggested. Again, an inherent assumption is that C-C bond scission happens in all metals from the same intermediate. While this may not be valid, it shows the potential of using these ML models for catalyst screening.

4. Conclusions

Accurate prediction of activation energies is crucial for evaluating the kinetic performance of catalysts. In this study, we compiled a dataset of 380 DFT-calculated activation barriers for hydrogenolysis reactions on various transition metals. We developed two methods to estimate activation energies. The first is an extension of the traditional BEP relationship, applied here to larger hydrocarbon species. Overall, BEP relationships still hold true, with dehydrogenation reactions being more accurately predicted than C-C cracking reactions. This finding aligns with the extent of structural change during the reaction and corroborates findings from previous studies [46,57,60]. These correlations are more accurate on Pt than Ru and Rh. However, when aggregating all data, BEP estimations result in relatively large errors.

To address shortcomings of traditional BEPs, we developed ML models using easily computable descriptors for the reactant, metal surfaces, and reaction type. The XGBoost model emerged as the most effective, demonstrating a mean absolute error of 0.19 eV without the reaction energy and 0.16 eV when including the reaction energy as a descriptor. Critical features include the homologous series (C-C vs. C-H reaction) and the metal atom's electronegativity and atomic mass. We also assessed the model's transferability, showing reasonable applicability in predicting activation barriers across transition metals. These advancements could speed up catalyst screening.

Finally, we showcased the use of the ML model to understand the impact of hydrocarbon chain length and branching on Ru, Rh, and Pt hydrogenation catalysts and provide insights into weak points where scission can happen, and coke or methane can form. Our model provides some of the first mechanistic insights into the impact of branching in polyolefins, including isotactic polypropylene and high vs. low density polyethylene, on the turnover frequency. We also demonstrated that these ML models can be valuable in catalyst screening. More work is needed to make these models more quantitative for polymer recycling. Future directions should include development and comparison of kinetic models to experimental data, accounting for polymer conformations and finite catalyst particle size, and refinement of reaction steps and parameters using DFT. The hydrogenolysis models developed here for single metals could easily be extended to bimetallic surfaces and bifunctional hydrocracking catalysts.

CRediT authorship contribution statement

Dionisios G Vlachos: Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Xue Zong:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **TJ (Tianjun) Xie:** Writing – original draft, Methodology, Data curation.

Author contributions

X.Z. and **T.X** performed the DFT calculations. **X.Z.** performed data analysis and developed the machine learning models. **D.G.V.** contributed to the idea and provided supervision. All authors contributed to analyzing the results and writing.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Data availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Acknowledgements

This work was supported as part of the Catalysis Center for Energy Innovation, an Energy Frontier Research Center funded by the US Dept. of Energy, Office of Science, Office of Basic Energy Sciences under award number DE-SC0001004. The earlier dataset was supported by the Department of Energy's Office of Energy Efficient and Renewable Energy's Advanced Manufacturing Office under Award Number DE-EE0007888-9.5. The Delaware Energy Institute gratefully acknowledges the support and partnership of the State of Delaware toward the RAPID projects.

Appendix A. Supporting information

The following information is available free of charge. (1) BEP correlations on each metal for C-C cracking and C-H breaking reactions; (2) Examples of transition state configurations for C-C and C-H breaking reactions; (3) Performance comparison of various tested machine learning (ML) models; (4) Summary of optimal hyperparameters for each ML model; (5) Summary of ML model performances; (6) Summary of ML model transferability performances. DFT calculation results for C-C cracking and C-H breaking reactions are provided in Supplementary_Data.xlsx file. Supplementary data associated with this article can be found in the online version at doi:10.1016/j.apcatb.2024.124070.

References

(1)P.A. Kots, T. Xie, B.C. Vance, C.M. Quinn, M.D. de Mello, J.A. Boscoboinik, C. Wang, P. Kumar, E.A. Stach, N.S. Marinkovic, L. Ma, S.N. Ehrlich, D.G. Vlachos, Electronic modulation of metal-support interactions improves polypropylene hydrogenolysis over ruthenium catalysts, Nat. Commun. 13 (1) (2022) 5186.
(2)A.L. Patrício Silva, J.C. Prata, T.R. Walker, A.C. Duarte, W. Ouyang, D. Barcelò, T. Rocha-Santos, Increased plastic pollution due to COVID-19 Pandemic: challenges and recommendations, Chem. Eng. J. 405 (2021) 126683.
(3)J. Zheng, S. Suh, Strategies to reduce the global carbon footprint of plastics, Nat. Clim. Chang. 9 (5) (2019) 374–378.
(4)J.E. Rorrer, C. Troyano-Valls, G.T. Beckham, Y. Román-Leshkov, Hydrogenolysis of polypropylene and mixed polyolefin plastic waste over Ru/C to produce liquid alkanes, ACS Sus. Chem. Eng. 9 (35) (2021) 11661–11666.
(5)C. Wang, K. Yu, B. Sheludko, T. Xie, P.A. Kots, B.C. Vance, P. Kumar, E.A. Stach, W. Zheng, D.G. Vlachos, A general strategy and a consolidated mechanism for low-methane hydrogenolysis of polyethylene over ruthenium, Appl. Catal. B Environ.. 319 (2022) 121899.
(6)X. Chen, Y. Wang, L. Zhang, Recent progress in the chemical upcycling of plastic wastes, ChemSusChem 14 (19) (2021) 4137–4151.
(7)A. Rahimi, J.M. García, Chemical recycling of waste plastics for new materials production, Nat. Rev. Chem. 1 (6) (2017) 1–11.
(8)L.D. Ellis, N.A. Rorrer, K.P. Sullivan, M. Otto, J.E. McGeehan, Y. Román-Leshkov, N. Wierckx, G.T. Beckham, Chemical and biological catalysis for plastics recycling and upcycling, Nat. Catal. 4 (7) (2021) 539–556.
(9)P.A. Kots, B.C. Vance, D.G. Vlachos, Polyolefin plastic waste hydroconversion to fuels, lubricants, and waxes: a comparative study, React. Chem. Eng. 7 (1) (2021) 41–54.
(10)P.A. Kots, S. Liu, B.C. Vance, C. Wang, J.D. Sheehan, D.G. Vlachos, Polypropylene plastic waste conversion to lubricants over Ru/TiO2 catalysts, ACS Catal. 11 (13) (2021) 8104–8115.
(11)S.D. Jaydev, A.J. Martín, J. Pérez-Ramírez, Direct conversion of polypropylene into liquid hydrocarbons on carbon-supported platinum catalysts, ChemSusChem 14 (23) (2021) 5179–5185.
(12)G. Celik, R.M. Kennedy, R.A. Hackler, M. Ferrandon, A. Tennakoon, S. Patnaik, A. M. LaPointe, S.C. Ammal, A. Heyden, F.A. Perras, M. Pruski, S.L. Scott, K. R. Poeppelmeier, A.D. Sadow, M. Delferro, Upcycling single-use polyethylene into high-quality liquid products, ACS Cent. Sci. 5 (11) (2019) 1795–1803.
(13)F. Zhang, M. Zeng, R.D. Yappert, J. Sun, Y.-H. Lee, A.M. LaPointe, B. Peters, M. M. Abu-Omar, S.L. Scott, polyethylene upcycling to long-chain alkylaromatics by tandem hydrogenolysis/aromatization, Science 370 (6515) (2020) 437–441.
(14)J.E. Rorrer, G.T. Beckham, Y. Román-Leshkov, Conversion of polyolefin waste to liquid alkanes with ru-based catalysts under mild conditions, JACS Au 1 (1) (2021) 8–12.
(15)Y. Nakaji, M. Tamura, S. Miyaoka, S. Kumagai, M. Tanji, Y. Nakagawa, T. Yoshioka, K. Tomishige, Low-temperature catalytic upgrading of waste polyolefinic plastics into liquid fuels and waxes, Appl. Catal. B Environ. 285 (2021) 119805.
(16)M. Salciccioli, M. Stamatakis, S. Caratzoulas, D.G. Vlachos, A review of multiscale modeling of metal-catalyzed reactions: mechanism development for complexity and emergent behavior, Chem. Eng. Sci. 66 (19) (2011) 4319–4355.
(17)J.E. Sutton, D.G. Vlachos, Building large microkinetic models with first-principles′ accuracy at reduced computational cost. Chem. Eng. Sci. 121 (2015) 190–199.
(18)A.H. Motagamwala, J.A. Dumesic, MicrokinetiC Modeling: A Tool for Rational Catalyst Design, Chem. Rev. 121 (2) (2021) 1049–1076.
(19)G.R. Wittreich, K. Alexopoulos, D.G. Vlachos, Microkinetic modeling of surface catalysis, in: W. Andreoni, S. Yip (Eds.), in: Handbook of Materials Modeling, Springer International Publishing, Cham, 2020, pp. 1377–1404.
(20)A.H. Motagamwala, M.R. Ball, J.A. Dumesic, Microkinetic analysis and scaling relations for catalyst design, Annu. Rev. Chem. Biomol. Eng. 9 (1) (2018) 413–450.
(21)K. Reuter, Ab initio thermodynamics and first-principles microkinetics for surface catalysis, Catal. Lett. 146 (3) (2016) 541–563.
(22)B.W.J. Chen, L. Xu, M. Mavrikakis, Computational methods in heterogeneous catalysis, Chem. Rev. 121 (2) (2021) 1007–1048.
(23)M.G. Evans, M. Polanyi, Inertia and driving force of chemical reactions, Trans. Faraday Soc. 34 (0) (1938) 11–24.
(24)J.N. Bronsted, Acid and basic catalysis, Chem. Rev. 5 (3) (1928) 231–338.
(25)R.P. Bell, C.N. Hinshelwood, The theory of reactions involving proton transfers, Proc. R. Soc. Lond. A 154 (882) (1997) 414–429.
(26)V. Pallassana, M. Neurock, Electronic factors governing ethylene hydrogenation and dehydrogenation activity of pseudomorphic PdML/Re(0001), PdML/Ru(0001), Pd(111), and PdML/Au(111) surfaces, J. Catal. 191 (2) (2000) 301–317.
(27)A. Michaelides, Z.-P. Liu, C.J. Zhang, A. Alavi, D.A. King, P. Hu, Identification of general linear relationships between activation energies and enthalpy changes for dissociation reactions at surfaces, J. Am. Chem. Soc. 125 (13) (2003) 3704–3705.
(28)M. Salciccioli, D.G. Vlachos, Kinetic modeling of Pt catalyzed and computation-driven catalyst discovery for ethylene glycol decomposition, ACS Catal. 1 (10) (2011) 1246–1256.
(29)S. Choi, Y. Kim, J.W. Kim, Z. Kim, W.Y. Kim, Feasibility of activation energy prediction of gas-phase reactions by machine learning, Chem. – A Eur. J. 24 (47) (2018) 12354–12358.
(30)J. Xu, X.-M. Cao, P. Hu, Perspective on computational reaction prediction using machine learning methods in heterogeneous catalysis, Phys. Chem. Chem. Phys. 23 (19) (2021) 11155–11179.
(31)G.H. Gu, C. Choi, Y. Lee, A.B. Situmorang, J. Noh, Y.-H. Kim, Y. Jung, Progress in computational and machine-learning methods for heterogeneous small-molecule activation, Adv. Mat. 32 (35) (2020) 1907865.
(32)C.A. Grambow, L. Pattanaik, W.H. Green, Deep learning of activation energies, J. Phys. Chem. Lett. 11 (8) (2020) 2992–2997.
(33)B. Wang, T. Gu, Y. Lu, B. Yang, Prediction of energies for reaction intermediates and transition states on catalyst surfaces using graph-based machine learning models, Mol. Catal. 498 (2020) 111266.
(34)A.R. Singh, B.A. Rohr, J.A. Gauthier, J.K. Nørskov, Predicting chemical reaction barriers with a machine learning model, Catal. Lett. 149 (9) (2019) 2347–2354.
(35)F. Göltl, M. Mavrikakis, Generalized brønsted-evans-polanyi relationships for reactions on metal surfaces from machine learning, ChemCatChem 14 (24) (2022) e202201108.
(36)D.J. Hutton, K.E. Cordes, C. Michel, F. Göltl, Machine learning-based prediction of activation energies for chemical reactions on metal surfaces, J. Chem. Inf. Model. *63* (19) (2023) 6006–6013.
(37)T. Xie, G.R. Wittreich, D.G. Vlachos, Multiscale modeling of hydrogenolysis of ethane and propane on Ru(0001): implications for plastics recycling, Appl. Catal. B: Env. *316* (2022) 121597.
(38)G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B 54 (16) (1996) 11169–11186.
(39)G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Comp. Mat. Sci. 6 (1) (1996) 15–50.
(40)A.H. Larsen, J.J. Mortensen, J. Blomqvist, I.E. Castelli, R. Christensen, M. Dulak, J. Friis, M.N. Groves, B. Hammer, C. Hargus, E.D. Hermes, P.C. Jennings, P.B. Jensen, J. Kermode, J.R. Kitchin, E.L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K.S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, The atomic simulation environment—a python library for working with atoms, J. Phys.: Condens. Matter 29 (27) (2017) 273002.
(41)G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B 59 (3) (1999) 1758–1775.
(42)G. Kresse, J. Hafner, Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements, J. Phys.: Condens. Matter 6 (40) (1994) 8245.
(43)J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865–3868.
(44)S. Grimme, S. Ehrlich, L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, J. Comput. Chem. 32 (7) (2011) 1456–1465.
(45)H.J. Monkhorst, J.D. Pack, Special points for Brillouin-Zone integrations, Phys. Rev. B 13 (12) (1976) 5188–5192.
(46)J.E. Sutton, D.G. Vlachos, A theoretical and computational analysis of linear free energy relations for the estimation of activation energies, ACS Catal. 2 (8) (2012) 1624–1634.
(47)G. Henkelman, B.P. Uberuaga, H. Jónsson, A climbing image nudged elastic band method for finding saddle points and minimum energy paths, J. Chem. Phys. 113 (22) (2000) 9901–9904.
(48)G. Henkelman, H. Jónsson, A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives, J. Chem. Phys. 111 (15) (1999) 7010–7022.
(49)G.H. Vineyard, Frequency factors and isotope effects in solid state rate processes, J. Phys. Chem. Sol. 3 (1) (1957) 121–127.
(50)F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (85) (2011) 2825–2830.
(51)J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (5) (2001) 1189–1232.
(52)J.H. Friedman, Stochastic gradient boosting, Comput. Stat. Data. Anal. 38 (4) (2002) 367–378.
(53)S. Lundberg, S.-I. Lee, A Unified Approach to Interpreting Model Predictions (2017).
(54)M.L. Waskom, Seaborn: statistical data visualization, J. Open. Source Softw. 6 (60) (2021) 3021.
(55)J. Bergstra, D. Yamins, D. Cox, Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures, in: Proceedings of the Thirtieth International Conference on Machine Learning; PMLRR (2013) 115–123.
(56)J. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for hyper-parameter optimization. in: Advances in Neural Information Processing Systems Vol. 24, Curran Associates, Inc, 2011.
(57)J.E. Sutton, P. Panagiotopoulou, X.E. Verykios, D.G. Vlachos, Combined DFT, microkinetic, and experimental study of ethanol steam reforming on Pt, J. Phys. Chem. C 117 (9) (2013) 4691–4706.
(58)M.-L. Yang, Y.-A. Zhu, C. Fan, Z.-J. Sui, D. Chen, X.-G. Zhou, DFT study of propane dehydrogenation on pt catalyst: effects of step Sites, Phys. Chem. Chem. Phys. 13 (8) (2011) 3257–3267.
(59)H.-F. Wang, Z.-P. Liu, Comprehensive mechanism and structure-sensitivity of ethanol oxidation on platinum: new transition-state searching method for resolving the complex reaction network, J. Am. Chem. Soc. 130 (33) (2008) 10996–11004.
(60)R.A. van Santen, M. Neurock, S.G. Shetty, Reactivity theory of transition-metal surfaces: a brønsted−evans−polanyi linear activation energy−free-energy analysis, Chem. Rev. 110 (4) (2010) 2005–2048.
(61)M. García-Mota, B. Bridier, J. Pérez-Ramírez, N. López, Interplay between carbon monoxide, hydrides, and carbides in selective alkyne hydrogenation on palladium, J. Catal. 273 (2) (2010) 92–102.
(62)R. García-Muelas, N. López, Statistical learning goes beyond the d-band model providing the thermochemistry of adsorbates on transition metals, Nat. Commun. 10 (1) (2019) 4687.
(63)Sinfelt, J.H. Specificity in Catalytic Hydrogenolysis by Metals, in: Advances in Catalysis; Eley, D.D., Pines, H., Weisz, P.B., Eds.; Academic Press, 1973; 23, 91–119.
(64)J.L. Carter, J.A. Cusumano, J.H. Sinfelt, Hydrogenolysis of *n*-Heptane over Unsupported Metals, J. Catal. 20 (2) (1971) 223–229.

(65)E. Kikuchi, M. Tsurumi, Y. Morita, Hydrogenolysis and isomerization of n-pentane on group viii transition metals, J. Catal. 22 (2) (1971) 226–236.
(66)K. Takahashi, I. Miyazato, Rapid estimation of activation energy in heterogeneous catalytic reactions via machine learning, J. Comput. Chem. 39 (28) (2018) 2405–2408.
(67)X. Zong, D.G. Vlachos, Exploring structure-sensitive relations for small species adsorption using machine learning, J. Chem. Inf. Model. 62 (18) (2022) 4361–4368.
(68)Mozer, M.C.; Jordan, M.I.; Petsche, T. Advances in neural information processing systems 9, in: Proceedings of the 1996 Conference; MIT Press, 1997.
(69)L. Breiman, Random Forests, Mach. Learn. 45 (1) (2001) 5–32.
(70)S.M. Lundberg, G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, S.-I. Lee, From local explanations to global understanding with explainable AI for trees, Nat. Mach. Intell. 2 (1) (2020) 56–67.
(71)D.W. Flaherty, D.D. Hibbitts, E. Iglesia, Metal-catalyzed C–C bond cleavage in alkanes: effects of methyl substitution on transition-state structures and stability. J. Am. Chem. Soc. 136 (27) (2014) 9664–9676.
(72)D.D. Hibbitts, D.W. Flaherty, E. Iglesia, Role of branching on the rate and mechanism of C–C cleavage in alkanes on metal surfaces. ACS Catal. 6 (1) (2016) 469–482.
(73)D.D. Hibbitts, D.W. Flaherty, E. Iglesia, Effects of chain length on the mechanism and rates of metal-catalyzed hydrogenolysis of n-alkanes, J. Phys. Chem. C. 120 (15) (2016) 8125–8138.
(74)A. Almithn, D. Hibbitts, Comparing rate and mechanism of ethane hydrogenolysis on transition-metal catalysts, J. Phys. Chem. C. 123 (9) (2019) 5421–5432.
(75)J.H. Sinfelt, D.J.C. Yates, Catalytic hydrogenolysis of ethane over the noble metals of group VIII, J. Catal. 8 (1) (1967) 82–90.